# Genetic analysis and molecular mapping of an *Rf* gene from *Helianthus angustifolius* for a new cytoplasmic male-sterile line

**Zhao Liu[1], Dexing Wang[2], Chao-Chien Jan[3]**

[1]Department of Plant Sciences, North Dakota State University, Fargo, ND 58102, USA, Email: zhao.liu@ars.usda.gov
[2]Liaoning Academy of Agricultural Sciences, Shenyang, 110161, China, Email: wang--dexing@163.com
[3]USDA-ARS, Northern Crop Science Laboratory, Fargo, ND 58102, USA,
Email: chaochien.jan@ars.usda.gov

## ABSTRACT

- The combination of cytoplasmic male-sterile (CMS) and the corresponding fertility restoration (*Rf*) genes is a critical tool in large-scale hybrid seed production of sunflower. A new CMS line 514A, derived from *H. tuberosus*, was obtained from a scientific exchange with the Liaoning Academy of Agricultural Sciences, Liaoning, China. In order to identify its *Rf* genes, 33 maintainer and restorer lines from five countries for the CMS PET-1 cytoplasm and 20 tester lines from the USDA-ARS-NCSL were used, but none of them restored fertility for CMS 514A. The aims of this study were: (1) to identify the *Rf* gene from different sources; (2) to analyze the genetics of the *Rf* gene; and (3) to locate the *Rf* gene on a molecular linkage map.

- Crossing and backcrossing methods were used for detecting the *Rf* genes and reducing the chromosome number to 2n=34. Genomic *in situ* hybridization (GISH) was used to verify the alien chromosome or segments. Two $F_2$ populations were used for mapping the gene. The $F_3$ or testcross progeny were used to phenotype the $F_2$ individuals. Bulk segregation analysis was used to screen the 370 SSR markers spanning the 17 sunflower linkage groups. Additional SSR and EST-SSR markers from the candidate linkage group (LG) were employed for mapping the gene. The deviation analyses of the fertility trait and marker loci were compared with the expected Mendelian ratios in the $F_2$ generation using the Chi-square test. The MAPMAKER/Exp version 3.0b software was used for linkage analysis, with the Kosambi mapping function.

- Five interspecific tetraploid amphiploids and one hexaploid *Helianthus* accession were crossed to CMS 514A, with the $F_1$ progenies pollinated by one amphiploid (Amp *H. angustifolius/*P 21, 2n=68) that produced all male-fertile plants with 2n=51. After several backcrosses with HA 89, male-fertile plants with 2n=35 were obtained. Genetic study of two male-fertile plants (2n=35) suggested that the *Rf* gene was located on the alien chromosome. The fertile plants with 2n=34 were obtained after backcrossing. The alien chromosome or segments were detected by GISH. Molecular mapping located this *Rf* gene on LG 3 of the public sunflower SSR map. Eight markers were linked to this gene, covering a genetic distance of 19.8 and 9.7 cM in the two mapping populations, respectively, with the closest marker ORS13 at a distance of 2.8 cM.

- The *Rf* gene was identified in the amphiploid of *H. angustifolius/*P 21. This *Rf* gene was expected to be on one alien chromosome in the fertile progenies with 2n=35. The *Rf* gene has been transferred into the progenies with 2n=34 using the backcross method, and was mapped to LG 3 of the sunflower map. The closely linked marker can be used for marker-assisted selection for this *Rf* gene. The study provides a new CMS/*Rf* gene system. Further analysis of the alien segment in the fertile $F_2$ individuals will be conducted using the GISH technique. Meiotic analysis will also be used for chromosome pairing study in the future.

- A single CMS PET-1 originating from wild *H. petiolaris* subsp. *petiolaris* Nutt. has been widely used for commercial sunflower hybrid seed production since it was discovered in1969. The maintainer and restorer lines for CMS PET-1 failed to restore the fertility of CMS 514A, suggesting this CMS is different than CMS PET-1. The *Rf* gene was finally detected in an amphiploid of *H. angustifolius/*P 21, and was located on LG 3. The results provide an alternative CMS/*Rf* gene system which will expand the diversity of sunflower germplasm for breeding, and help to understand the interaction mechanism between the cytoplasm and nuclear genes.

**Key words:** cytoplasmic male-sterility - fertility restoration gene - molecular mapping

## INTRODUCTION

The combination of cytoplasmic male-sterile (CMS) and the corresponding fertility restoration (*Rf*) genes is a critical tool in large-scale hybrid seed production of cultivated sunflower (*Helianthus annuus* L.) (Horn et al., 2003; Serieys, 1996). For over 40 years, the commercial sunflower hybrid seed industry has depended on a single CMS, CMS PET-1, which originated from wild *H. petiolaris* subsp. *petiolaris* Nutt., discovered in 1969 (Horn et al., 2003; Leclerq, 1969). The $Rf_1$ gene is considered responsible for fertility restoration of this CMS (Dominguez-Gimenez and Fick, 1975; Miller and Fick, 1997; Jan and Vick, 2007). Alternative CMS/*Rf* gene systems will expand the diversity of sunflower germplasm for breeding, reduce the risk of using a single CMS/*Rf* system, and help understand the interaction mechanism between nuclear and cytoplasmic genes. Seventy-two sunflower CMS sources have been identified up to 2005 (Serieys, 2005), but only about half have available corresponding *Rf* genes, and one to four dominant *Rf* genes are necessary for fertility restoration depending on the materials (Serieys, 1996). However, only six *Rf* genes have been mapped on sunflower genetic maps, including $Rf_1$, $Msc_1$, $Rf_3$-RHA 340 and $Rf_3$-RHA 280 for CMS PET-1, $Rf_4$ for a new alloplasmic CMS GIG2, and *Rf-PEF1* for CMS PEF1. These *Rf* genes have been mapped to different positions on the public sunflower maps (reviewed by Liu et al., 2011).

A CMS line, CMS 514A, was derived from *H. tuberosus* at the Liaoning Academy of Agricultural Sciences, Liaoning, China. Thirty-three maintainer and restorer lines from five countries used with the CMS PET-1 cytoplasm failed to restore fertility in CMS 514A. The objectives of this study were: (1) to identify the *Rf* gene from different sources; (2) to analyze the genetics of the *Rf* gene; and (3) to locate the *Rf* gene on a molecular linkage map.

## MATERIALS AND METHODS
### Plant materials
Twenty tester lines from the USDA-ARS-NCSL, five tetraploid interspecific amphiploids, and one hexaploid *Helianthus* accession were used to cross with CMS 514A. The fertile $F_1$ plants were backcrossed with HA 89 to obtain male-fertile (MF) plants with 2n=35 and 34 chromosomes. $BC_4F_1$ (2n=35) and $BC_4F_2$ (2n=34) progeny from CMS 514A x *H. angustifolius*/P 21 amphiploid (2n=68) were used for cytogenetic analysis. Two $F_2$ mapping populations derived from two $F_1$ individuals of CMS 514A/6/[(CMS 514A//Amp of *H. angustifolius*/P 21/3/2*HA 89/4/HA 821/5/HA 89) SIB] were used for the mapping of the *Rf* gene. They are referred to as G08 and G09 hereafter.

### Pollen fertility analysis and $F_2$ genotype confirmation
A standard Alexander's procedure was employed to check pollen fertility of each MF plant (Alexander 1969) using an Axioplan2 Imaging microscope (Zeiss, Germany). Three fields of view with 200-400 pollen grains were analyzed. The percentage of the fertile pollen grains of the total was used for analysis. The fertile $F_2$ individuals were self-pollinated to obtain $F_3$ seeds. The plants with poor pollen fertility in the G09 population were crossed with HA 89 in order to obtain seeds for progeny test. The $F_3$ and testcross progeny test of the two $F_2$ populations was carried out with 20-50 plants from each $F_2$ plant in the field at Fargo, ND, and Casselton, ND, in 2009.

### DNA extraction
Genomic DNA was extracted according to the protocol of the Qiagen DNAeasy 96 Plant Kit (Qiagen, Valencia, CA, USA). The bulked segregant analysis method was used for polymorphism screening (Michelmore et al., 1991), with equal quantities of DNA from 10 plants for each bulk. Four bulks were used for polymorphism screening, including the homozygous fertile and sterile bulks of $F_2$ plants with 2n=34, and a fertile bulk with 2n=35 and a sterile bulk with 2n=34 in the $BC_4F_2$ progeny.

### Molecular marker screening and PCR analysis
In total, 370 pairs of SSR primers mapped to the 17 sunflower linkage groups from the Compositae database (http://compositdb.ucdavis.edu) were used for polymorphism screening between the four bulks. An additional 73 SSR markers and 28 expressed sequence tag (EST)-SSR markers from the candidate LG 3 from 23 maps in the Sunflower CMap Database (http://sunflower.uga.edu/cgi-bin/cmap/map search) were used to screen the polymorphism among the parents and the $F_1$ plants. Polymorphic markers were used for genotyping of the mapping population after confirmation. The PCR amplification and genotyping data was conducted following Liu et al. (2011).

**Statistical analysis and linkage map construction**

The deviation analyses of the fertility trait and marker loci were compared with the expected Mendelian ratios in the $F_2$ generation using the Chi-square test. The MAPMAKER/Exp version 3.0b program (Whitehead Institute, Cambridge, MA, USA) (Lander et al., 1987) was used for linkage analysis of the phenotypes and molecular genotypes, following Liu et al. (2011).

## RESULTS

**The *Rf* gene was identified in an amphiploid of *H. angustifolius*/P 21**

$F_1$ plants from the crosses of CMS 514A with 20 cultivated lines, one accession of *H. californicus*, and four of the five amphiploids were all male-sterile (MS), including the cross with P 21 (Table 1). However, the crosses of CMS 514A with amphiploids of *H. angustifolius*/P 21 were all MF (2n=47-51), indicating the *Rf* gene was from *H. angustifolius*.

**Table 1.** The $F_1$ progeny fertility restoration of 20 tester lines, five amphiploids (Amp) and one hexaploid *H. californicus* accession crossed with CMS 514A

| Cross | Plant No. | Fertile Plant No. |
|---|---|---|
| CMS 514A/HA 821 | 44 | 0 |
| CMS 514A/HA 801 | 21 | 0 |
| CMS 514A/Senaca | 18 | 0 |
| CMS 514A/Hopi Dye | 30 | 0 |
| CMS 514A/Smena | 38 | 0 |
| CMS 514A/Luch | 37 | 0 |
| CMS 514A/Issanka | 42 | 0 |
| CMS 514A/RHA 266 | 42 | 0 |
| CMS 514A/RHA 376 | 36 | 0 |
| CMS 514A/RHA 274 | 32 | 0 |
| CMS 514A/RHA 294 | 21 | 0 |
| CMS 514A/Armavir | 39 | 0 |
| CMS 514A/VNIIMK | 37 | 0 |
| CMS 514A/P 21 | 32 | 0 |
| CMS 514A/RCMSG 3 | 46 | 0 |
| CMS 514A/RHA 290 | 47 | 0 |
| CMS 514A/HA 89 | 36 | 0 |
| CMS 514A/RCMG 1 | 39 | 0 |
| CMS 514A/RCMG 2 | 47 | 0 |
| CMS 514A/Peredovik | 46 | 0 |
| CMS 514A//*H. californicus*/HA 89 | 4 | 0 |
| CMS 514A//Amp of *H. atrorubens*/P 21 | 6 | 0 |
| CMS 514A//Amp of *H. mollis*/P 21 | 11 | 0 |
| CMS 514A//Amp of *H. cusickii*/P 21 | 6 | 0 |
| CMS 514A//Amp of *H. grosseserratus*/P 21 | 10 | 0 |
| CMS 514A//Amp of *H. angustifolius*/P 21 | 13 | 13 |
| Total | 780 | |

**The *Rf* gene was integrated into the cultivated sunflower**

The $F_1$ plants were backcrossed several times with HA 89 in order to reduce the chromosome number to 2n=34. Among the 36 $BC_1F_1$ plants, 11 were MF, with 2n=38-46. Eight out of 34 $BC_3F_1$ plants were MF, but all with 2n=35. Two plants with 2n=34 and nine with 2n=35 were obtained in the 69 $BC_4F_1$ plants. Testcross progenies of CMS plants pollinated with two MF $BC_4F_1$ plants with 2n=35 chromosomes resulted in 55 MS plants (2n=34) and four MF plants with 2n=35, suggesting that the *Rf* gene was located on the extra chromosome. Self-pollination of eight 2n=35 MF plants resulted in one MF progeny with 2n=36, 10 with 2n=35, and two with 2n=34, and the remaining 49 were MS. The results indicated that the chromosome recombination rate between the chromosomes of *H. angustifolius* and the cultivated sunflower was very low.

**Fertility segregation in the mapping populations**

The G08 $F_2$ population was composed of 204 individuals, including 17 homozygous MF, 60 heterozygous MF, and 88 MS plants, with 39 plants not having enough seeds for $F_3$ progeny testing. The G09 $F_2$ population included 245 individuals, with 16 homozygous MF, 110 heterozygous MF, and 86 MS plants identified in 212 $F_3$ or testcross progeny individuals. Chi-square test indicated the homozygous MF: heterozygous MF: homozygous MS phenotypes severely deviated from the 1:2:1 ratio ($\chi^2 \geq 46$, P=0.00) (Table 2).

In the two $F_2$ mapping populations, a large variation in fertility was observed (15-100%), with an average of 79% and 85% for the two populations, respectively. In the G08 population, about 85% of the plants had fertility above 50%, while 65% had fertility above 80%. A similar tendency was observed in the G09 $F_2$ population, with about 82% of the plants having fertility above 50%, and 73% with fertility above 80%.

**Table 2.** Segregation of the *Rf* locus in the two $F_2$ populations (G08 and G09) derived from the cross CMS 514A /6/ [(CMS 514A//Amp of *H. angustifolius*/P 21/3/2*HA 89/4/HA 821/5/HA 89) SIB]

| Population | Number of $F_2$ plants[a] | Observed number[b] | | | | Ratio tested | $\chi^2$ | P |
|---|---|---|---|---|---|---|---|---|
| | | A | H | B | C | | | |
| G08 | 204 | 88 | | | 116 | 1:3 | 35.79 | 0.00 |
| G08 | 165 | 88 | 60 | 17 | | 1:2:1 | 73.38 | 0.00 |
| G09 | 245 | 86 | | | 159 | 1:3 | 13.33 | 0.00 |
| G09 | 212 | 86 | 110 | 16 | | 1:2:1 | 46.53 | 0.00 |

a: For G08 and G09 $F_2$ populations , 39 and 33 plants were fertile, but no $F_3$ or testcross progeny were tested because of limited seed.

b: Symbols: A, homozygous MS (*rfrf*); H, heterozygous MF(*Rfrf*); B, homozygous MF (*RfRf*); C, *RfRf* or *Rfrf*.
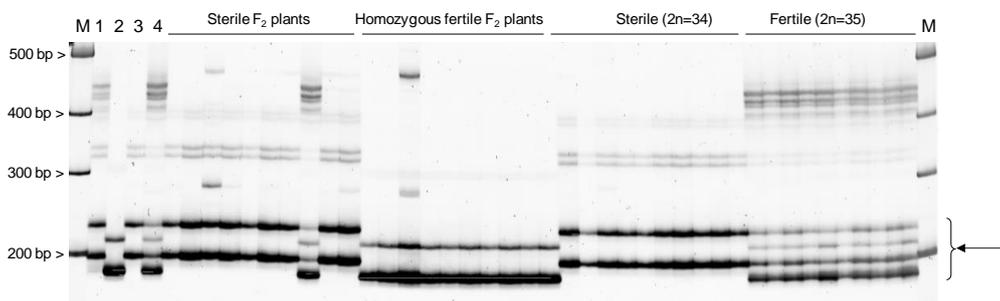


**Fig. 1.** Representative results of the ORS488 marker among the four bulks and the individuals constituting each bulk, respectively, on a non-denaturing polyacrylamide gel.
1: Sterile $F_2$ bulk; 2: Homozygous fertile $F_2$ bulk; 3: Sterile bulk (2n=34); 4: Fertile bulk (2n=35); The bulks 3 and 4 consist of sterile plants (2n=34) and the fertile plants (2n=35) from the self progenies of a $BC_4F_1$ fertile plant (2n=35), respectively; M indicates 100-bp plus ladder Gelpilot (Qiagen); the arrow indicates the co-dominant marker.

**The *Rf* gene was located on LG 3**

Bulked segregant analysis identified four polymorphic markers using the 370 pairs of SSR primers mapped to the 17 sunflower linkage groups among the four bulks (2n=34 and 35), including ORS433, ORS488, ORS822, and ORS1114. The polymorphic markers were confirmed using the 10 individuals constituting each of the four bulks. Representative confirmation results of the markers are shown in Figure 1, using the ORS488 primer. ORS488 and ORS1114 were mapped to LG 3 of Tang et al. (2003), whereas ORS433 and ORS822 were multi-loci markers, which were mapped to LG 10 and LG 1, besides LG 3. This *Rf* gene was proposed to be on LG 3. An additional 73 SSR and 28 EST-SSR markers on LG 3 from 23 maps in the Sunflower CMap Database (http://sunflower.uga.edu/cgi-bin/cmap/map search) were used to screen the polymorphism among the parents and the $F_1$ plants. Another two SSR markers, ORS13 and ORS525, and two EST-SSR markers, HT088 and HT734, were obtained. After confirmation of the polymorphism using the 10 individuals constituting each of the four bulks, these markers were used

for genotyping of the mapping populations. A map was constructed for each of the mapping populations, using the eight markers mentioned above. The suppressed recombination was observed for the two maps, covering a genetic distance of 19.8 and 9.7 cM, respectively. The closest marker to this $Rf$ gene was ORS13, which was at a distance of 2.8 cM.

## DISCUSSION

Alternative CMS/$Rf$ system will provide more diversity and reduce the risk of using the single CMS PET-1/$Rf_1$ system for commercial sunflower hybrid seed production. We reported the identification of an $Rf$ gene for a new source of CMS. In total, 59 different sources were used to identify the $Rf$ gene for CMS 514A. The $Rf$ gene was only detected in an amphiploid of $H.\ angustifolius$/P 21. These results suggest that this $Rf$ gene is different than all other $Rf$ genes reported so far. Compared to the average 98% restored pollen fertility of the $Rf_1$ gene (Hybrid 894 with the pedigree CMS HA 89 x RHA 274, Seiler, 2000), and that of the $Rf_3$ gene from RHA 280 at 94% (Liu et al., 2011), the average pollen fertility of the $Rf$ gene identified here was about 80%, with a large variation in both $F_2$ populations. The $Rf$ gene was mapped to LG 3 of the sunflower SSR public map (Tang et al., 2003), with eight linked markers. A severe deviation was detected for both the $Rf$ gene and the linked markers in both mapping populations. Additionally, low rate of transmission of the alien chromosome or segments into the cultivated sunflower was detected during backcrossing. Meiotic GISH analysis of the MF plants with 2n=35, as well as on the MF plants in the $F_2$ mapping populations may help to explain the chromosome pairing and recombination from a cytogenetic perspective.

Six $Rf$ genes have been previously mapped on sunflower genetic maps, with only the $Rf_4$ gene reported on LG 3 (Feng and Jan, 2008; reviewed by Liu et al., 2011). The closest marker linked to the $Rf_4$ gene is ORS1114, which is at a distance of 0.9 cM. The $Rf$ gene reported here was also mapped to LG 3, with the closest marker ORS13 2.8 cM away from it. Considering the suppressed recombination of the linkage group, the actual distance between ORS13 and the $Rf$ gene may be even larger than the distance on this map. Therefore, the $Rf$ gene for CMS 514A and the $Rf_4$ gene for CMS GIG2 are probably not the same gene. Their allelic relationship still needs to be determined. The combination of the markers closely linked to the $Rf$ genes will be helpful for marker-assisted selection during sunflower breeding. The CMS/$Rf$ system can also be used to study the interaction of the nuclear gene and the cytoplasm in sunflower. The DNA rearrangement and genes responsible for CMS in sunflower have been studied in the mitochondrial genome from different sources. Extensive efforts to create a saturated map of $Rf_1$ have been made in order to isolate this gene using a map-based cloning method (Horn et al., 2003; Kusterer et al., 2005). Further study of the diversity of the mitochondrial genes and even isolation of the $Rf$ genes will help to reveal the fertility restoration mechanism in sunflower, and to study the evolution of the $Rf$ genes in different plants.

## REFERENCES

Alexander. M.P. 1969. Differential staining of aborted and nonaborted pollen. Stain Technol. 44:117-122.

Dominguez-Gimenez, J., and G.N. Fick. 1975. Fertility restoration of male-sterile cytoplasm in wild sunflowers. Crop Sci. 15:724-726.

Feng, J., and C.C. Jan. 2008. Introgression and molecular tagging of $Rf_4$, a new male fertility restoration gene from wild sunflower $Helianthus\ maximiliani$ L. Theor. Appl. Genet. 117:241-249.

Horn, R., B. Kusterer, E. Lazarescu, M. Prüfe, and W. Friedt. 2003. Molecular mapping of the $Rf1$ gene restoring pollen fertility in PET1-based $F_1$ hybrids in sunflower ($Helianthus\ annuus$ L.). Theor. Appl. Genet. 106:599-606.

Jan, C.C., and B.A. Vick. 2007. Inheritance and allelic relationships of fertility restoration genes for seven new sources of male-sterile cytoplasm in sunflower. Plant Breed. 126:213-217.

Kusterer, B., R. Horn, and W. Friedt. 2005. Molecular mapping of the fertility restoration locus *Rf1* in sunflower and development of diagnostic markers for the restorer gene. Euphytica 143:35-42.

Lander, E.S., P. Green, J. Abrahamson, A. Barlow, M.J. Daly, S.E. Lincoln, and L.A. Newberg. 1987. MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics 1:174-181.

Leclercq, P. 1969. Une stérilité male cytoplasmic chez le tournesol. Ann. Amélior Plant 19:99-106.

Liu, Z., S. Mulpuri, J. Feng, B.A. Vick, and C.C. Jan. 2011. Molecular mapping of the $Rf_3$ fertility restoration gene to facilitate its utilization in breeding confection sunflower. Mol. Breed. doi: 10.1007/s11032-011-9563-0.

Michelmore, R.W., I. Paran, and R.V. Kesseli. 1991. Identification of markers linked to disease resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. Proc. Natl. Acad. Sci. U.S.A. 88:9828-9832.

Miller, J.F., and G.N. Fick. 1997. Genetics of sunflower. p. 441-495. In: A.A. Schneiter (ed.), Sunflower Technology and Production. Agronomy Monograph 35. ASA, CSSA and SSSA, Madison, WI, USA.

Seiler, G.J. 2000. Registration of ten interspecific germplasms derived from wild perennial sunflower. Crop Sci. 40:587-588.

Serieys, H. 1996. Identification, study and utilization in breeding programs of new CMS sources. FAO Progress Report. Helia 19 (special issue):144-158.

Serieys, H. 2005. Identification, study and utilization in breeding programs of new CMS sources, in the FAO Subnetwork. p. 47-53. In: Proc. 2005 Sunflower Subnetwork Progress Report, FAO, Rome, Italy.

Tang, S., V.K. Kishore, and S.J. Knapp. 2003. PCR-multiplexes for a genome-wide framework of simple sequence repeat marker loci in cultivated sunflower. Theor. Appl. Genet. 107:6-19.